

# Лекция 2. Развитие и современное состояние технологий Ethernet

## 2.1 Некоторые характеристики «изначального интернета»

### Понятие «изначального Ethernet»

В 1983 г. был разработан стандарт 10Base5 семейства 802.3 для толстого коаксиального кабеля, в 1985 г. - стандарт 10Base2 для тонкого коаксиального кабеля, и в 1990 г. — стандарт 10Base-T для кабеля типа «витая» пара. Эти три стандарта будем условно называть «изначальным шинным Ethernet-ом», в отличие от разработанного позднее множества других стандартов, существенно развивших изначальную технологию Ethernet и выведших ее на лидирующие позиции в большинстве «экологических ниш» традиционного применения других сетевых технологий.

Поясним значение упомянутых названий: 10Base5, 10Base2 и 10Base-T. Используемые во всех этих названиях число «10» и слово «BASE» означают соответственно пропускную способность 10 Мбит/сек и широкополосный (baseband) характер использования среды передачи кабеля. Числа 5 и 2 означают количество десятых долей дюйма (0,24 см), соответствующих толщине (диаметру) коаксиального кабеля, а буква «Т» используется для обозначения кабеля типа «витая пара».

### Структура кадра Ethernet

Передаваемый через сеть кадр Ethernet включает следующие поля (они входят в кадр в указанном порядке): преамбула, признак начала кадра, заголовок кадра, данные (пакет протокола 3-го уровня) и поле CRC.

**Преамбула** имеет длину 7 байт (56 бит) и битовое значение из чередующихся нулей и единиц: «10101010 . . . 10».

**Признак начала кадра** имеет длину 1 байт (8 бит) и значение «1010101011».

**Заголовок** кадра имеет длину 14 байт, его структура рассматривается ниже.

**Данными является пакет** протокола более высокого (3-го) уровня, его длина не должна превышать 1500 байт и не может быть меньше 46 байт. Если реальная длина пакета 3-го уровня меньше 46, этот пакет перед его инкапсуляцией в кадр дополняется до 46 байт.

Поле **CRC (Cyclic Redundancy Code)** — циклический код избыточности) имеет длину 4-байтного слова и формируется в процессе передачи поля данных путем циклического суммирования последовательности слов поля данных (с переносом возникающих разрядов переполнения для суммирования значений младшего разряда).

Отметим, что поля преамбулы и признака начала кадра входят в состав кадра, передаваемого через среду передачи, но отбрасываются принимающей стороной. Их длина не входит в длину принятого кадра. Эти поля предназначены для синхронизации получателя кадра с его отправителем и определения «точки начала собственно кадра».

Существует **четыре стандарта заголовка кадра Ethernet**, отличающихся использованием двух последних байт заголовка и наличием возможных дополнительных полей в заголовке. В наши задачи не входит рассмотрение всех этих стандартов, рассмотрим лишь наиболее распространенный из них — стандарт **Ethernet Type II**. В этом стандарте заголовок кадра включает 3 поля, следующих в указанном порядке: MAC-адрес получателя, MAC-адрес отправителя и поле длина/тип (см. рис. 1).

|           |              |                  |                   |           |        |     |
|-----------|--------------|------------------|-------------------|-----------|--------|-----|
| 7         | 1            | 6                | 6                 | 2         | 0-1500 | 4   |
| преамбула | начало кадра | адрес получателя | адрес отправителя | длина/тип | данные | CRC |

Рис. 1. Формат кадра Ethernet Type II. Вверху указаны размеры полей в байтах

**Длина полей MAC-адресов**, естественно, составляет по 6 байт и название этих полей само говорит о занесенных в них значениях. Двухбайтное поле **длина/тип** используется либо для указания кода протокола более высокого уровня (например, коды протоколов IP, ICMP, ARP), в рамках которого пересылается настоящий кадр, либо для указания значения длины кадра. Какое именно из двух этих значений реально передается в заголовке кадра определяется по простому правилу. Поскольку общая длина кадра **не может превышать 1518** (14 байт заголовка + 1500 байт данных + 4 байта CRC), в качестве кодов протоколов более высокого уровня используются значения существенно превосходящие 1518 (в следующем параграфе вы узнаете, что в начало кадра могут добавляться дополнительные поля). Эти коды используются для определения протокола более высокого уровня, модуль реализации которого надо передать принятый в кадре пакет. В этом случае длина кадра просто подсчитывается автоматически (аппаратурой сетевой карты).

Отметим, что **максимальное значение длины принимаемого кадра называется термином MTU (Maximum Transfer Unit — максимальная единица передачи)**. В шинных сетях Ethernet значение MTU равно 1518. **В коммутируемых сетях Ethernet** при использовании специальных протоколов, расширяющих функциональные возможности этих сетей, **требуемое значение MTU может быть больше**.

### **Метод CSMA/CD (МДКН/ОС)**

Множественный доступ с контролем носителя и обнаружением столкновений (**шинная топология**)

### **Применение витой пары для построение шинной сети Ethernet**

Витая пара (twisted pair — закрученная пара) представляет собой несколько пар скрученных изолированных проводников, заключенных в общую изолирующую оболочку, как это показано на рис. 2. Чаще всего встречаются кабели с четырьмя и двумя парами.

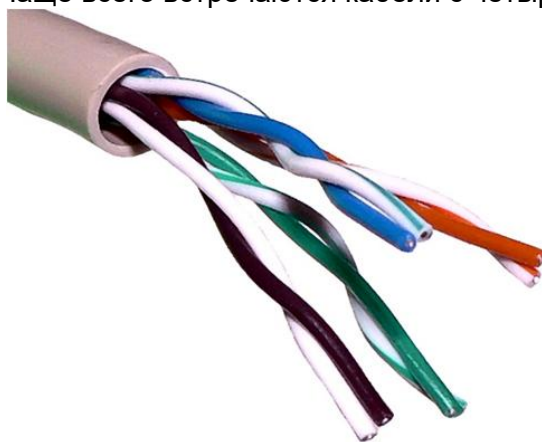


Рис. 2. Кабель витая пара

Завитие проводников осуществляется с целью уменьшения электромагнитных помех от внешних источников, например, от кабелей питания, а также для уменьшения взаимного влияния передаваемых по парам сигналов. В отличие от плоского кабеля токи, наводимые в витках пары электромагнитным полем внешних помех, нейтрализуют друг друга. Чтобы электромагнитное поле одной пары не влияло на соседние, их завивают с разным шагом.

Каждый из компьютеров сегмента подключается к хабу или коммутатору отрезком кабеля длиной не больше 100 метров, с установленными на его концах коннекторами RJ-45 (см. рис. 3). Один из этих коннекторов вставляется в порт (разъем RJ-45) хаба или коммутатора, другой — в разъем RJ-45 сетевой карты компьютера.



Рис. 3. Коннектор RJ-45

Для обжимки коннекторов витой пары понадобится специальный инструмент (см. рис.4), который вдавливает ножи коннектора внутрь жил кабеля. Зачищать жилы кабеля перед обжимкой не требуется. Достаточно вставить провода в правильном порядке в гнезда коннектора, наложить обжимные клещи и сжать рукоятки.

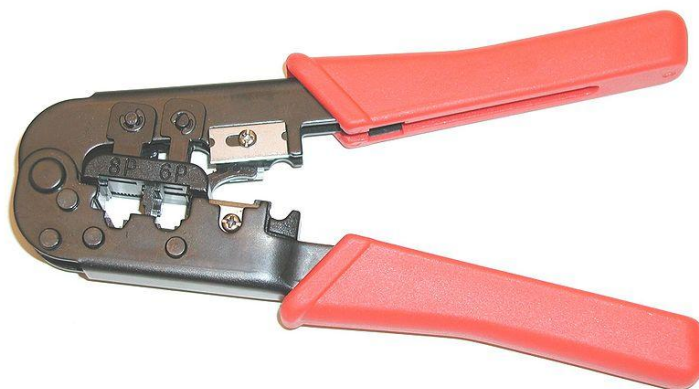


Рис. 2.12. Обжимные клещи (кримпер)

Из-за исторических причин, связанных с совместимостью с другими кабелями, существует две схемы обжимки витой пары, описанные в стандарте TIA/EIA-568: А и В (см. рис.4). Автору чаще всего приходилось сталкиваться со схемой В, однако современное оборудование отлично работает как для кабелей, обжатых по обоим стандартам А и В, так и для так называемых кросс-кабелей с одним концом, обжатым по стандарту А, а другим — по В.

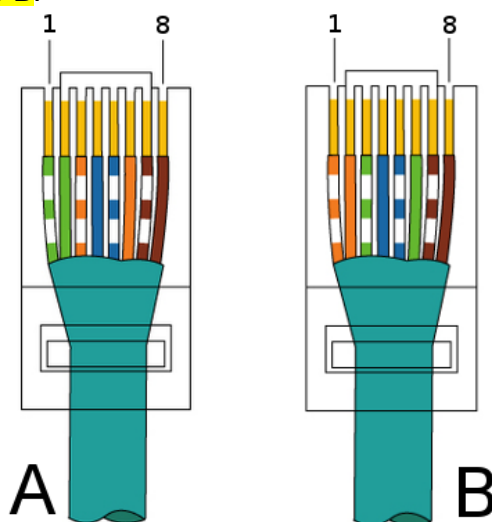


Рис. 4. Схемы обжимки витой пары TIA/EIA-568: А и В

### Организация электропитания устройств по витой паре (Power over Ethernet - PoE)

В наиболее распространенной сегодня витой паре категории 5е при передаче сигнала по технологии Fast Ethernet (100Base-T, 100Мбит/с) используются всего две пары жил (проводов) кабеля из четырех. Для схемы TIA/EIA-568В это: оранжевая (жилы 1 и 2)

и зеленая (жилы 3 и 6). Поэтому в продаже можно часто встретить кабель с всего двумя витыми парами.

Неиспользуемые для передачи сигнала жилы в 4х-парном кабеле можно использовать либо для подачи электропитания к подключаемому маломощному устройству (например – к IP-телефону). Ограничение по силе тока для двух пар — не более 400 мА, по напряжению — 48 В. Для длинных кабелей будет наблюдаться падение напряжения. Для тока 400мА оно составит приблизительно 4 вольта на 100 м кабеля. При подаче питания на сетевые устройства по витой паре следует уделить внимание металлу, из которого сделаны жилы кабеля. Дорогую медь иногда заменяют омедненным алюминием, который имеет большее сопротивление и меньшую допустимую силу тока. На передаче сигнала это не сказывается, но подать электропитание той же мощности по алюминиевому кабелю уже не получится.

Организовать питание по витой паре довольно несложно «вручную», но есть коммутаторы, которые одновременно с функцией передачи сигнала предоставляют дополнительную возможность подачи электроэнергии. Данная технология называется Power over Ethernet (PoE). Она позволяет автоматически настраивать требуемое для устройства напряжение и организовывать так называемое фантомное питание — передачу электроэнергии по сигнальным проводам. Настройка напряжения в PoE производится по сложному протоколу, производящему серию проб подачи различных напряжений и измерений силы тока. Поэтому сетевое оборудование, поддерживающее стандарт PoE стоит на порядок дороже обычных аналогов. Чтобы сэкономить, многие производители пренебрегают протоколом согласования питающего напряжения и продают свои сетевые устройства вместе с комплектом адаптеров, подающих фиксированное напряжение на жилы витой пары. Это так называемый пассивный PoE (Passive PoE). Он не совместим со стандартным PoE, описанным в IEEE 802.3af-2003 и IEEE 802.3at-2009. Использование несовместимых технологий может привести к сгоранию оборудования.

## 2.2 Обзор высокоскоростных стандартов Ethernet

Начиная с середины 1990-х годов разработчиками технологий семейства Ethernet с завидной методичностью выпускались все более скоростные стандарты этой технологии, первые из которых имели варианты, рассчитанные на применение различных типов кабеля (витой пары и волоконно-оптического кабеля). Более поздние (и более скоростные стандарты) ориентированы на применение только волоконно-оптических кабелей.

Первой скоростной технологией семейства Ethernet стала технология Fast Ethernet (стандарты 100Base-T для витой пары категории 5 и 100Base-F для волоконно-оптического кабеля), разработанные в 1995 году и обеспечивающие передачу данных с пропускной способностью 100 Мбит/сек.

Затем были созданы технологии Gigabit Ethernet (стандарты 1000Base-TX и 1000Base-FX, разработанные в 1998 году), 10Gigabit Ethernet (стандарты 10GBase-xx, где xx - одно из множеств сочетаний символов) и 40Gigabit Ethernet (стандарты 40GBase-xx), разработанные с 2003 года, 100Gigabit Ethernet (стандарты 100GBase-xx, разработанные с 2010 года). Цифры, указанные в названиях этих технологий и стандартов говорят сами за себя. Полный перечень стандартов технологий семейства Ethernet с характеристиками используемых типов кабелей приведен, например, в <https://habrahabr.ru/post/208202/>.

К настоящему времени разработаны стандарты группы Terabit Ethernet: 200Gigabit Ethernet и 400Gigabit Ethernet. И опять вероятно, что на этом рассмотренный процесс дальнейшего «ускорения» технологии Ethernet не остановится.

## 2.3. Назначение, принципы работы и классификация мостов и коммутаторов Ethernet

**Мосты Ethernet (bridge)** - коммуникационные устройства, предназначенные для объединения в один сегмент нескольких сегментов сети, были впервые разработаны в конце 1990-х годов. **Вначале мосты были двухпортовыми** и обеспечивали соединение 2-х коаксиальных сегментов Ethernet (витой пары в ту пору еще не было) для увеличения вдвое длины сегмента. Такие мосты называются также повторителями (**repeater**). Затем были созданы мосты, предназначенные для объединения сегментов, построенных с использованием разных технологий и многопортовые (с 4-мя и более портами) мосты. Примером моста с разнотипными портами является **мост-модем Ethernet**, имеющий порты (разъемы) RJ-10 и RJ-45, предназначенные для подключения соответственно к телефонной линии и сети Ethernet.

Известны следующие типы мостов: прозрачные, транслирующие и инкапсулирующие:

- **Прозрачные мосты** имеют порты, работающие по одной и той же технологии и эффект их использования практически “не виден” пользователям подключенных к сегментам компьютеров. О наличии таких мостов пользователи могут судить лишь по большой длине объединенного сегмента.

- **Транслирующие мосты** имеют 2 порта разных типов и обеспечивают объединение сегментов, работающих по различным технологиям, выполняя трансляцию формата передаваемых кадров из формата отправителя в формат получателя. Наличие таких мостов заметно по различию типа кабельных систем, и сетевых карт и другого коммуникационного оборудования в различных частях объединенного сегмента. Примером такого моста является упомянутый выше мост-модем.

**Инкапсулирующие мосты** имеют 2 порта различных типов используются для передачи через сеть одной технологии (назовем ее несущей) кадров другой технологии (назовем ее пассажиром). Для организации такой передачи при вхождении кадров технологии-пассажира в сеть несущей технологии через инкапсулирующий мост эти кадры инкапсулируются (помещаются в поле данных) несущей технологии, при выходе такого кадра из несущей сети через инкапсулирующий мост того же типа выполняется декапсуляция инкапсулированного кадра-пассажира и он передается в целевую сеть, построенную на основе соответствующей технологии.

А теперь, отметим один **принципиально важный момент в организации работы многопортовых мостов**. Для повышения уровня масштабируемости создаваемых с их использованием “больших” сегментов эти **коммутаторы могут выполнять одновременную пересылку данных между любыми парами их портов**. Это обеспечило **повышение уровня масштабируемости** таких сегментов во столько раз, сколько пар портов могут работать одновременно. А это число равно целой части от половины количества портов.

Рассмотренное свойство оказалось настолько привлекательным, что оно было использовано для создания нового типа коммуникационных устройств, названных коммутаторами (switch) Ethernet. **Коммутатор Ethernet** - это многопортовый (обычно с количеством портов от 5 до 128 и более) с портами, работающими на основе технологий одного семейства (Ethernet), конструктивно выполненный в виде хаба. При достаточно большом числе портов (от 16-24 без ограничения сверху) коммутаторы обычно оформляются в виде, предназначенном для монтажа в коммуникационные стойки и имеют высоту 1 или несколько юнитов (в зависимости от количества портов). Отметим, что порты одного коммутатора Ethernet могут “работать” с разной скоростью и обеспечивать подключение к разнотипным средам передачи (витой паре и волоконно-оптическому кабелю). Но **формат кадра, применяемых различными технологиями семейства Ethernet унифицирован и это принципиально важно**. И также важно отметить, что в рамках каждого соединения коммутаторы обеспечивают передачу данных с максимально возможной для участвующих в соединениях портов коммутаторов скоростью (естественно, что в случае различия максимальных скоростей работы указанных портов, коммутатором будет обеспечиваться пересылка данных со скоростью более медленного порта).

**На базе коммутаторов могут строиться иерархические структуры (деревья или, что значит то же самое - иерархические звезды)**. В **корне** должен находиться **наиболее скоростной коммутатор** (через него проходят все потоки данных между компьютерами,

подключенными к различным поддеревьям. С удалением коммутатора от корня требования к его максимальным скоростным характеристикам снижаются. В рассмотренной структуре можно полностью отказаться от использования шинных сегментов. В этом случае проблемы масштабирования по количеству компьютеров для построенного в виде иерархии коммутаторов сегмента Ethernet просто не существует.

## 2.4 Алгоритм работы коммутаторов, некоторые вопросы их реализации, обзор дополнительных функций управляемых коммутаторов

### Алгоритм работы многопортовых мостов и коммутаторов

Поскольку коммутаторы Ethernet являются всего лишь разновидностью (хотя и очень важной и очень широко востребованной) прозрачных многопортовых мостов Ethernet, коммутаторы и мосты применяют один и тот же алгоритм перенаправления кадров, который далее мы будем называть алгоритмом работы коммутатора. Этот алгоритм включает 2 основных алгоритма: алгоритм коммутации и алгоритм самообучения коммутатора, который является лишь небольшим шагом первого алгоритма. Работа алгоритма коммутации основана на использовании таблиц коммутации, существующих для каждого из портов коммутаторов. В таблице коммутации каждого из портов содержатся MAC-адреса получателей, которые доступны через этот порт (возможно транзитно, через один или несколько промежуточных коммутаторов). Вначале таблица коммутации пуста, затем в процессе обучения эта таблица заполняется.

Рассмотрим действия коммутатора, выполняемые при получении кадра через некоторый порт. Вначале коммутатор извлекает из кадра MAC-адрес отправителя и, если его еще нет в таблице этого порта, то заносит туда этот адрес. Систематическое выполнение этого действия каждым портом коммутатора полностью обеспечивает обучение коммутатора, ведь каждый подключенный через любой порт коммутатора рано или поздно пошлет в сеть какой-то кадр либо для обращения к некоторому другому компьютеру сети, либо в ответ на такое обращение.

После выполнения этого “обучающего” шага алгоритма коммутатор проверяет наличие MAC-адреса получателя кадра в таблицах всех портов коммутатора, включая порт из которого прибыл этот кадр. Если адрес получателя кадра содержится в таблице этого (входного) порта, кадр просто игнорируется, ибо всем получателям из шинного сегмента, подключенного к указанному порту этот кадр передавался одновременно с передачей его коммутатору.

Если адрес получателя находится в таблице какого-то еще порта, кадр отправляется в этот порт.

Если же адреса получателя нет ни в одной таблице MAC-адресов для всех портов коммутатора, то принятый кадр широкоэвещательно рассылается во все порты, кроме того, из которого он был принят. Тогда если получатель с указанным MAC-адресом действительно подключен к сегменту сети, в который входит рассматриваемый коммутатор, этот кадр будет доставлен получателю, возможно через ряд транзитных коммутаторов. Если же получателя с указанным MAC-адресом в сегменте просто нет, кадр будет просто проигнорирован всеми компьютерами и маршрутизаторами, подключенными коммутаторам самых нижних уровней иерархии в древовидной структуре коммутаторов (листьев в дереве коммутаторов) либо непосредственно, либо через шинные подсегменты.

Таким образом, при выполнении рассмотренного алгоритма всеми коммутаторами древовидной структуры (единичный коммутатор - частный случай древовидной структуры, состоящей из одного корня) все пересылаемые через эту структуру кадры корректно доставляются по назначению.

Дополнительно отметим, что в рассмотренном алгоритме работы коммутатора при взаимодействии компьютеров с уже известными коммутатору MAC-адресами трафик этих компьютеров попадает только в “подсегменты” источника и получателя кадров. В “подсегменты”, подключенные к сторонним портам коммутатора попадает только широкоэвещательный трафик, вызванный передачей первого кадра от еще неизвестного

коммутатору компьютера. Поэтому возможности “прослушки” трафика, обусловленные соответствующими возможностями шинных сегментов Ethernet, существенно уменьшаются при соединении шинных сегментов через коммутаторы и практически сводятся “на нет” при подключении всех компьютеров сети непосредственно к портам коммутаторов.

### **Некоторые вопросы внутреннего устройства коммутатора и особенностей работы коммутаторов.**

Коммутатор имеет достаточно сложное внутреннее устройство, включающее, как минимум:

- процессор с памятью и специализированной операционной системой (ОС)
- порты (с возможными очередями к ним, размещенные в памяти коммутатора)
- устройство параллельной передачи данных между портами.

Процессор с памятью и ОС необходимы для программной реализации алгоритма коммутации и ряда других обязательных и/или дополнительных функций коммутатора, рассматриваемых ниже после рассмотрения общих вопросов организации работы коммутатора..

#### ***Методы борьбы с перегрузкой портов.***

Выше упоминалось, что темп поступления данных в некоторые порты коммутатора эпизодически может быть выше максимальной скорости передачи данных через этот порт. В этом случае кадры, поступающие в моменты занятости порта передачей некоторого пакета, заносятся в одну из двух очередей порта: входную или выходную. **Во входную очередь** порта заносятся кадры, пришедшие в него извне через подключенную к порту кабельную систему. Такое занесение во входную очередь выполняется в том случае, если коммутатор не успевает выполнить перенаправление кадра до поступления во входной порт следующего кадра. **В выходную очередь** порта заносятся кадры, направляемые в этот порт из других портов коммутатора и предназначенные для передачу вовне через кабельную систему, подключенную к этому порту.

И входные и выходные очереди могут переполняться. В этом случае вновь поступающие в очередь кадры выбрасываются. Между тем, очереди к порту коммутатора могут не успевать отправлять находящиеся в их выходных очередях пакеты из-за того, что со стороны кабельной системы в этот порт передается слишком “плотный” поток кадров, не оставляющий коммутатору времени, требуемого отправки всех кадров из очереди.

Для обеспечения возможности передачи портом информации вовне вне зависимости от уровня входящей нагрузки на порт коммутатора применяются специальные **методы борьбы с перегрузками портов**, позволяющие сократить размер очередей к коммутаторов и уменьшить уровень потерь предназначенных к отправке через этот порт кадров. К числу этих методов относятся метод агрессивного поведения коммутатора и метод обратного давления на порт коммутатора.

Суть **метода агрессивного поведения коммутатора** состоит в том, что в случае возникновения столкновений между кадрами, пересылаемыми из порта коммутатора и из некоторой сетевой карты сегмента, подключенного к этому порту искусственно повышается до 1 вероятность того, что коммутатор выйдет из режима ожидания раньше любой из сетевых карт. Делается это следующим образом. На самом деле время ожидания после столкновения является не просто случайной величиной включает вначале некоторую технологическую паузу  $t_r$ , а затем на паузу случайной величины  $t_c$ , так, что суммарное время задержки  $t_s = t_r + t_c$ . Аналогичная технологическая пауза выдерживается и между передачей указанными устройствами (коммутатором и сетевыми картами) нескольких кадров подряд.

Агрессивное поведение коммутатора состоит в том, что для него величина  $t_r$ , выбирается намного меньшей ( $\ll$ ), чем для сетевых карт. В результате эффект агрессивного поведения порта коммутатора состоит в том, что даже при интенсивном поступлении в этот порт пакетов из подключенной к нему кабельной системы

(подсегмента), он всегда будет успевать будет успевать отправлять в этот порт кадры, пересылаемые в него из других портов.

**Метод обратного давления на порт** состоит в том, что на чрезмерно активный порт, входные потоки данных которого чрезмерно загружают выходные очереди других портов, организуется отправка фиктивных пакетов (не адресованных никакому реальному получателю) с тем, чтобы путем захвата части пропускной способности канала снизить поступающую с него нагрузку на выходные порты коммутатора.

### ***Два возможных способа передачи пакетов через коммутатор.***

В рамках рассмотренного алгоритма работы коммутатора он может использовать для передачи пакетов (кадров) один из двух следующих способов: **способ сквозной передачи пакетов** или **способ передачи пакетов с буферизацией**.

При сквозной передаче пакета сразу после считывания из входного порта адреса его получателя и определения выходного порта организуется передача в этот выходной порт информации, поступающей из входного порта без ее промежуточной буферизации.

При передаче пакета с буферизацией он полностью считывается из входного порта в буферную память коммутатора.

При первом способе величина задержки при передаче пакета через коммутатор во много раз меньше, чем при втором способе. Однако этот способ имеет ряд недостатков. Наиболее значимыми из них являются два следующих. Во-первых, корректность кадра обнаруживается только по завершению его передачи. А это значит, что порою коммутатор занят транзитной передачей некорректных кадров, что создает дополнительную нагрузку на сеть. А, во-вторых, первый способ не в состоянии справиться с ситуацией, в которой темп поступления информации, предназначенной для передачи через некоторый выходной порт, хотя бы временно превышает пропускную способность этого порта. Это возможно, если указанная информация поступает из нескольких входных портов и/или из более скоростного порта. Этот недостаток настолько существенен, что на практике обычно используется передача пакетов с их промежуточной буферизацией.

### ***Разновидности устройств передачи данных между портами коммутатора.***

Устройство передачи данных между портами коммутатора должно обладать пропускной способностью, достаточной для одновременной передачей кадров между любыми парами портов этого коммутатора. Наиболее известными способами организации такого устройства являются коммутирующая матрица, высокоскоростная шина и специальная организация с многоходовой памятью.

**Коммутирующая матрица** размера  $N \times N$  (где  $N$  - количество портов коммутатора) - это матричная структура, строки которой подключены к входным портам коммутатора, а столбцы - к его выходным портам. На пересечении строк и столбцов матрицы находятся коммутирующие элементы, связанные со своими соседями и/или с портами коммутатора (если одним из соседей коммутирующего элемента является порт). Каждый коммутирующий элемент может быть в процессе коммутации переведен в одно из двух возможных положений: положение продолжения передачи кадра по строке (до достижения столбца выходного порта) и положение продолжения передачи кадра по столбцу (к выходному порту). Указанное переключение состоит в **переключении соответствующих электрических цепей**. Поэтому этот способ реализации устройства передачи данных **обеспечивает сквозную передачу кадров** (вспомните достоинства и недостатки такого способа передачи). Этот метод применим лишь в случае, если пропускные способности всех портов коммутатора одинаковы. Кроме того, для реализации этого способа требуется использование  $N^2$  переключательных элементов, что существенно увеличивает стоимость коммутаторов с довольно большим количеством портов. Поэтому с учетом этих недостатков, а также упомянутых выше недостатков сквозной передачи пакетов коммутирующие матрицы практически не используются при реализации коммутаторов Ethernet.

**Высокоскоростная шина** коммутатора обеспечивает случайный доступ к своей среде передачи данных и должна предусматривать в своей реализации средство "арбитража" конкурентного доступа, подобное методу CSMA/CD. И, поскольку в подобных средах



передачи при загрузке их пропускной способности более чем на 2/3 происходит резкий всплеск уровня загрузки пропускной способности трафиком повторной передачи пакетов после их попадания в столкновения, реальная пропускная способность высокоскоростной шины должна быть ориентировочно в полтора раза выше, чем сумма пропускных способностей любых соединений между парами портов коммутатора. Кроме того, возможные столкновения на этой высокоскоростной шине увеличивают задержки кадров при передаче их через коммутатор. Поэтому, несмотря на то, что рассматриваемый метод организации внутреннего коммуникационного устройства коммутатора является самым дешевым, он может использоваться на практике **лишь в самых дешевых моделях коммутаторов, не претендующих на достижение высоких показателей качества предоставляемого ими сервиса** быстрой (с минимальными задержками) передачи кадров.

**Специальная организация с многовходовой памятью** (допускающей одновременный доступ к этой памяти нескольких устройств) и параллельным подключением к этой памяти контроллеров портов коммутатора предусматривает независимо и параллельно выполняемые запись принятых контроллерами входных портов кадров в память и чтение контроллерами выходных портов тоже из памяти отправляемых кадров. Такой способ не создает никаких дополнительных задержек и одинаково хорошо работает при любых вариантах взаимной пропускной способности портов коммутатора. Поэтому несмотря на относительно высокую стоимость многовходовой памяти и контроллеров параллельного доступа к ней, рассмотренный способ применяется во всех коммутаторах, хоть сколь либо претендующих на обеспечение гарантированно низких задержек при коммутации кадров.

#### **Обзор дополнительных возможностей управляемых коммутаторов.**

**Если коммутатор предоставляет лишь рассмотренные выше возможности, он называется неуправляемым.**

**Управляемые коммутаторы предоставляют ряд дополнительных возможностей,** в число которых могут входить все или некоторые из следующих возможностей:

- Поддержка применения протокола SNMP
- Фильтрация кадров по их MAC-адресам
- Поддержка протокола STP (возможности резервирования каналов)
- Поддержка виртуальных локальных сетей VLAN
- Поддержка приоритизации кадров
- Некоторые другие возможности.

**Поддержка применения протокола SNMP** (Simple Network Management Protocol - простой протокол управления сетью), давшая название этому классу коммутаторов является обязательной. На всех управляемых коммутаторах установлено программное обеспечение агента SNMP, обеспечивающего возможности удаленного (через сеть) контроля состояния коммутатора и изменения его настроек

**Фильтрация кадров по их MAC-адресам** предназначена для обеспечения возможности взаимной изоляции некоторых компьютеров с целью обеспечения их взаимной информационной безопасности. Эта фильтрация состоит в выбрасывании кадров, удовлетворяющих условиям фильтрации, и может выполняться каждым портом коммутатора с целями “не выпускация” из некоторого “подсегмента” исходящих из него кадров и/или “не впускация” в “подсегмент” кадров, поступающих в него из других портов коммутации. Однако, поскольку условия фильтрации выражаются в терминах индивидуальных MAC-адресов и не позволяют легко (без серьезных усилий по настройке и поддержанию списков фильтрации) ограничить доступ для определенных подсетей. Поэтому указанными средствами фильтрации кадров практически не пользуются, предпочитая применение средств фильтрации IP-пакетов маршрутизаторами..

**Протокол STP** (Spanning Tree Protocol - протокол остового (от слова “остов”) дерева) предназначен для обеспечения возможности добавлять в древовидные структуры коммутаторов дополнительные каналы, соединяющие коммутаторы из разных

поддеревьев. Потребность использования таких каналов для резервирования путей доступа в случае отказа каналов древовидной иерархии очень высока. Поэтому протокол STP подробнее рассматривается далее.

**Виртуальные локальные сети VLAN** (Virtual LAN - виртуальная ЛВС) широко применяются для построения на базе единой структуры взаимосвязанных коммутаторов наложенных на эту структуру виртуальных (логических) сетей, объединяющих, например, компьютеры пользователей, входящих в определенную группу (лабораторию отдел и пр.). При построении VLAN достаточно широко используется протокол IEEE 802.1q. В виду высокой востребованности возможностей построения VLAN они рассматриваются нами на следующей лекции

Поддержка приоритезации кадров (для обеспечения различных уровней качества сетевого обслуживания) обеспечивается возможностями протокола IEEE 802.1p. Однако на этом уровне (уровне каналов данных) этими средствами пользуются редко, предпочитая соответствующие средства, основанными на использовании сетевых протоколов более высокого уровня (IP, MPLS).

## 2.5 Протокол STP поддержки применения резервных каналов в иерархических коммутируемых сетях

Стремление добавить в древовидную структуру коммутаторов каналы, соединяющие коммутаторы, находящиеся в различных поддеревьях, с целью обеспечения резервных путей передачи данных, которыми можно воспользоваться при отказе одного из каналов на соединяющем эти коммутаторы пути через древовидную структуру совершенно естественна.

Однако, если просто выполнить такое добавление резервного канала, то работоспособность сети будет нарушена из-за возникновения явления, называемого ширококестельным штормом. Ширококестельный шторм обязательно возникнет на каждом шаге обучения структуры коммутаторов, в случае, если в графе связи коммутаторов есть неориентированные циклы. А такой цикл обязательно возникает при соединении каналом любой пары коммутаторов из различных поддеревьев.

Проиллюстрируем возникновение ширококестельного шторма на примере структуры из трех коммутаторов: А, В и С, в которой корнем дерева является коммутатор А, а В и С - листья этого дерева, дополнительно соединенные резервным каналом между ними. Пусть через некоторый входной порт коммутатора А поступает кадр, MAC-адрес которого неизвестен ни одному из трех коммутаторов. Тогда коммутатор А ширококестельно отправит этот кадр коммутаторам В и С, см. Рис. 5а. Эти коммутаторы после получения кадра направят его друг другу, см. Рис. 5б. И после этого коммутаторы В и С по кадру коммутатору А, см. рис. 5в, после чего вновь повторится ситуация рис.2.9а и процесс начнет повторяться.

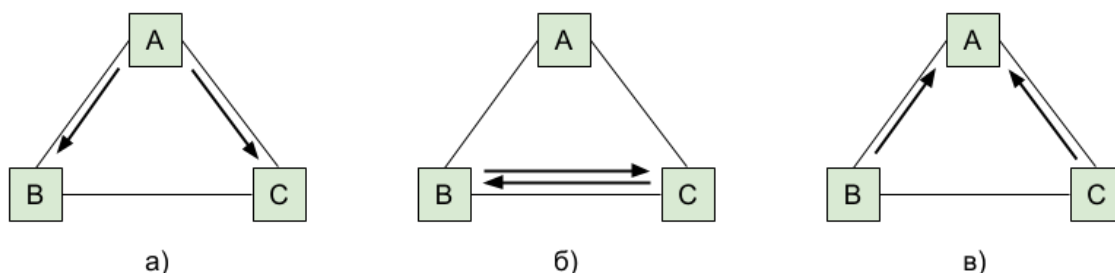


Рис. 5. Развитие ширококестельного шторма

Решение проблемы исключения ширококестельных штормов при использовании резервных каналов в дереве коммутаторов Ethernet обеспечивает протокол STP (Spanning Tree Protocol - протокол остового (от слова "остов") дерева).

Суть этого протокола состоит в том, что по коммуникационному графу сети коммутаторов строится покрывающее дерево (в терминах теории графов), называемое специалистами в области сетевых технологий остовым деревом. Это дерево обеспечивает связность всех вершин графа (коммутаторов) и, как любое дерево не

содержит циклических путей. Все дуги исходного коммуникационного графа, не вошедшие в покрывающий граф, помечаются как заблокированные и не могут использоваться для пересылки кадров. Таким образом исключается возможность возникновения широковещательного шторма. А при выходе из строя одного из каналов покрывающего дерева протокол STP строит новое покрывающее дерево, обеспечивая восстановление связности между всеми коммутаторами сети.

### **Общая организация работы протокола STP.**

В протоколе STP коммутаторы обмениваются друг с другом специальными пакетами BPDU (Bridge Protocol Data Unit - единица данных протокола мостов), в которые среди прочих полей входят 3 следующие, упоминаемые нами при рассмотрении алгоритма работы STP:

- ID корневого коммутатора
- ID коммутатора, отправившего данный BPDU
- Стоимость пути до корневого коммутатора

**ID-коммутатора** - это 8-байтное поле, включающее вложенные поля. Первое из этих полей имеет длину 2 байта и содержит **значение приоритета** (приоритет тем выше, чем меньше его значение), задаваемый администратором сети. А во втором 6-байтном поле содержится MAC-адрес коммутатора. Главным из этих вложенных полей является **приоритет, заданный администратором**. Дополнение его значением MAC-адреса обеспечивает уникальность каждого ID в случаях, когда администратор укажет для нескольких коммутаторов один и тот же приоритет. И в целом **8-байтное значение ID-коммутатора интерпретируется как его приоритет в своих претензиях на роль корневого коммутатора покрывающего дерева** и/или корневых коммутаторов поддеревьев этого дерева.

Очевидно, что поскольку **трафик между всеми поддеревьями корневого коммутатора проходит через корневой коммутатор**, скорость его работы должна быть не меньшей, чем скорость работы корневых коммутаторов этих поддеревьев. **Поэтому более скоростные коммутаторы должны получать более высокий приоритет**. Кроме того, при назначении приоритетов должны учитываться некоторые другие моменты. В частности, более высокий приоритет логично присваивать коммутаторам, установленным в точках сопряжения нескольких циклов. Это обеспечит быстроту обмена данными между поддеревьями, полученными при разрывании каждого из циклов.

**Стоимость пути до корня** - это сумма стоимостей портов, через которые надо пройти от корневого коммутатора к данному коммутатору (или от данного коммутатора к корневому). Стоимость порта пропорциональна времени прохождения через сеть кадра, отправленного из этого порта и скорости работы порта. Стоимость порта  $C$  (cost) может задаваться сетевым администратором, но обычно значение этой скорости вычисляется как отношение некоторой **эталонной пропускной способности (reference bandwidth -  $B_R$ )** к **пропускной способности порта ( $B$ )**:  $C = B_R / B$ . При этом в коммуникационном оборудовании компании Cisco (коммутаторах и маршрутизаторах)  $B_R = 10^8$ . Таким образом стоимость порта Fast Ethernet составит 1. Отметим, что в оборудовании других производителей значение  $B_R$  может быть другим, например,  $10^9$ .

### **Алгоритм работы протокола STP.**

Алгоритм работы протокола STP включает 2 последовательно выполняемых шага:

1. Выбор корневого коммутатора
2. Построение остового дерева с выбранным корнем

#### **Алгоритм выбора корневого коммутатора.**

Выбор корневого коммутатора выполняется по следующему алгоритму.

Каждый коммутатор широковещательно рассылает во все свои порты BPDU, в которых объявляет себя корневым.

После получения каждым из коммутаторов BPDU с ID меньшим собственного (остальные BPDU просто игнорируются), коммутатор перестает рекламировать себя и начинает рекламировать корневой коммутатор, указанный в принятом BPDU.

Порт, через который был принят этот BPDU, помечается как “корневой”, то есть ведущий к корневому коммутатору.

Рассмотренный процесс завершается через конечное число шагов (не превосходящее длину пути между самыми удаленными вершинами графа) выбором самого приоритетного корневого коммутатора.

#### **Алгоритм построения остового дерева.**

Коммутатор, выбранный корневым, отправляет во все свои порты BPDU с нулевой стоимостью пути до корня. Эти BPDU принимаются корневыми портами всех соседних коммутаторов. Каждый из таких соседей, получив BPDU от корневого коммутатора добавляет в стоимость пути до корня стоимость своего корневого порта и **рассылает обновленный BPDU во все порты, кроме корневого.**

Каждый из следующих коммутаторов при получении BPDU через несколько портов **выбирает для дальнейшей работы тот из них, который пришел по пути с минимальной стоимостью** (сумма стоимости пути до корня из BPDU и стоимости соответствующего корневого порта). **Остальные BPDU просто игнорируются, а порты, через которые пришли эти BPDU переводятся в заблокированное состояние, в котором они не используются для пересылки обычных пакетов, но полностью доступны для пересылки BPDU.** Потребность в такой пересылке может возникнуть при выходе из строя какого-то канала (или порта коммутатора) для начала построения нового остового дерева.

Дальнейшая работа с выбранным BPDU состоит в соответствующем увеличении стоимости пути до корня и в рассылке измененного BPDU во все порты, кроме корневого. Рассмотренный процесс завершится через конечное число шагов построением требуемого остового дерева.

Отметим, что действия, выполняемые на двух рассмотренных шагах алгоритма, каждый из которых совершает полный обход коммуникационного графа могут быть выполнены за один обход графа. Именно это делает протокол RSTP (Rapid STP - быстрый STP), название которого само говорит за себя.

#### **Пример применения резервных каналов в структурированной кабельной сети многоэтажного здания.**

Структурированные кабельные сети (СКС) - это кабельные системы, разработанные с использованием принципов и методов, обеспечивающих их высокую надежность и долговечность их эксплуатации. В соответствии с указанными принципами и методами СКС многоэтажного здания рекомендуется строить в соответствии со структурой, изображенной на Рис. 6.

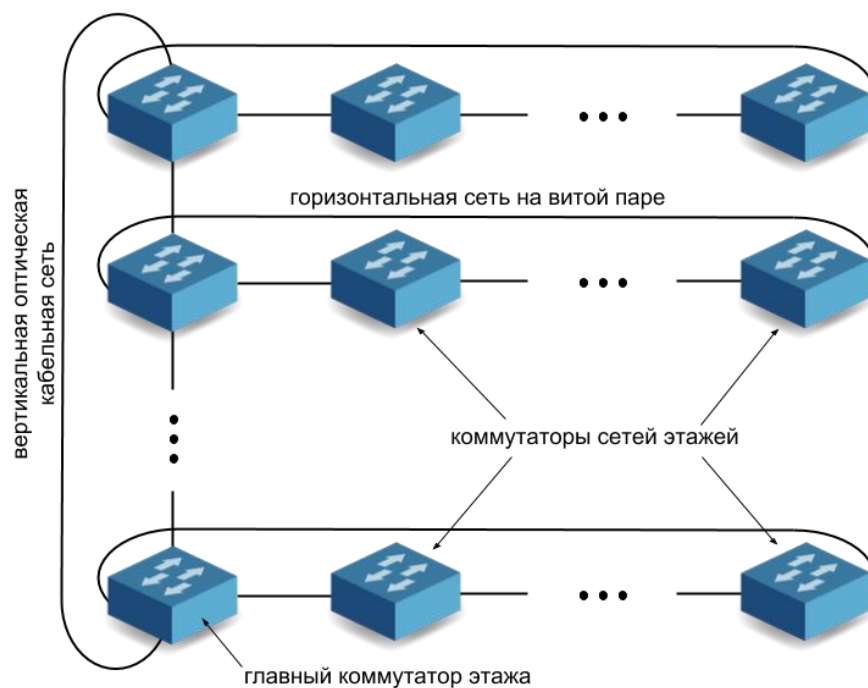


Рис. 6. Типовая организация СКС многоэтажного здания

В этой структуре отказ любого этажного коммутатора (кроме главного) работоспособность сети этажа (за исключением подсетей отказавшего коммутатора). Но выход из строя главного коммутатора этажа приведет к изоляции сети этого этажа. Поскольку к вертикальной кабельной сети предъявляются повышенные требования быстродействия и надежности и по ряду других соображений, рекомендуется строить ее на базе волоконно-оптического кабеля.